LAB ASSIGNMENT - 02

Name :- Aniket Kumar

Roll no :- 13

ERP :- 10321712O3

Batch :- 1

Problem Statement :-

Watson / NLTK Speech to Text and Text to Speech APIs - Using IBM watson Speech to Text service, convert an audio recording to text format Recording can be done on a mobile phone or using computer's microphone. Try with different audio formats like .mp3, .wav. Use smart formatting to convert numbers and currencies in proper text format. Use cases :-

- User dictates a paragh (speech), and it is converted to text.
- User utters a word, and the program spells it out (Speech → Text → Speech)
- Simple Calculator with voice commands (addition, subtraction, and square)

Objective :-

1. To study and explore IBM watson Speech to text and Text to Speech API's.
2. To learn concepts of Speech to Text and Text to Speech conversion in NLP.

Theory:- Explain following concepts.

- IBM Watson APIs.

IBM Watson is a AI platform which provides functionalities such as language Translation, visual Recognition, Data analysis, Speech to text and much more to application of clients. Every services require an IBM cloud account.

- Speech Recognition.

It is a field of Computer Science and Computational Linguistics that develop methodologies and technologies that enable the recognition and translation of spoken language into text by computers.

- Speech to Text & Text to Speech Conversion.

Speech to Text service helps to convert audio to subsequent text. The Speech to Text service by Watson API uses IBM's Speech recognition capabilities to produce transcripts of spoken audio. The service can transcribe speech from various languages and audio formats.

Text to Speech service helps to facilitate a textual data to be converted in subsequent language audio output. Here, the IBM Watson Text to Speech services helps in providing API that uses IBM's Speech Synthesis capabilities to synthesis text in natural-sounding speech in a variety of languages, dialets, and voices.

Algorithm / Implementation :-

1. Install all the dependencies and setup the microphone
2. Accept voice from the user with the mic
3. Remove noise and distortion from the Speech.
4. Convert the Speech or voice to text.
5. Now store the text as a string in variable
6. Print the string to confirm the equation.
7. Split the string into three parts: first operand, operator and second operand, and convert operand to integer.
8. Calculate the result.

Platform:- 64-bit Open Source Linux, IBM Watson Cloud, Json.

Input:- Voice Speech - any audio format file and text paragraph.

Output:- Audio file converted to Text and Text paragraph converted to Audio Signals (Speech)

Conclusion:- Hence, learned the concepts of Speech to Text and Text to Speech Conversion in NLP and implemented the same using IBM Watson API's.

FAQ's:-

1) Which Python modules are used in Speech recognition?

Ans:- Python modules used in Speed Recognition:-

1) apiai
2) assemblyai
3) google-cloud-Speech
4) pocketsphinx
5) Speech Recognition
6) watson-developer-cloud
7) wit

2) List IBM watson modules used for TTS and STT ?

Ans For Speech To Text, IBM watson provides:-

"Speech To Text-V1" from 'ibm_watson'

For Text to Speech, IBM watson provides:-

"Text To Speech-V1' from 'ibm-watson'.